FEBS Letters 548 (2003) 108–112

FEBS 27476

Differential binding of Sin3 interacting repressor domains to the PAH2 domain of Sin3A

Yuan-Ping Pang[a,b,c,d,*], Ganesh A. Kumar[a], Jin-San Zhang[e], Raul Urrutia[b,c,d,e,f,**]

[a] *Department of Molecular Pharmacology and Experimental Therapeutics, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*
[b] *Tumor Biology Program, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*
[c] *Molecular Neuroscience Program, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*
[d] *Mayo Clinic Cancer Center, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*
[e] *Gastroenterology Research Unit, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*
[f] *Department of Biochemistry and Molecular Biology, Mayo Foundation for Medical Education and Research, 200 First Street SW, Rochester, MN 55905, USA*

Received 16 April 2003; revised 23 June 2003; accepted 25 June 2003

First published online 7 July 2003

Edited by Takashi Gojobori

Abstract The Sin3 interacting domain (SID), originally described in the Mad family of repressors, is a novel transcriptional repressor domain that binds the PAH2 domain of corepressors Sin3A and Sin3B with high affinities. The conserved SID-like domains are reportedly present in five KLF proteins. However, the KLF SIDs and the Mad SIDs can be classified into two subtypes according to sequence similarity. Here, we report the finding from computational and experimental studies that the two subtypes of SID domains bind differentially to Sin3A. This finding offers insights into a mechanism of cell growth regulation by interactions of different subtypes of SID-containing repressor proteins with Sin3. It also provides the structural basis for developing selective modulators of Sin3.

Key words: Transcription; Zinc finger protein; KLF protein; SID; KLF11; Mad1

1. Introduction

The Sin3/HDAC corepressor complex is one of the well-characterized corepressor complexes and it is highly conserved from yeast to humans. The Sin3 protein functions as a large protein scaffold capable of supporting protein–protein interactions through its four repeats of ~100 residues known as paired amphipathic helix (PAH) domains. In addition to the intrinsic components of the complex including the HDAC1 and HDAC2, Sin3 associates with a wide variety of DNA-binding transcription factors using different PAH domains. Members of the Mad family of repressor proteins specifically interact with the second PAH domain (PAH2) of mSin3A through the Sin3 interacting domain (SID) [1]. The Mad1 SID contains at least 13 residues that form an amphipathic α-helix [2], as seen in the nuclear magnetic resonance (NMR) structures of the Mad1 SID complexed with the PAH2 domains [3,4]. A PAH2-interacting consensus sequence has accordingly been proposed to be 'ϕZZϕϕXAAXXϕ[E/D]', where X is any non-proline residue, ϕ is any bulky hydrophobic residue, and Z is any hydrophobic or polar/charged residue with a significant aliphatic component in the side chain [3]. The detailed structural information regarding the Mad1 SID complex is insightful in understanding structure and function of the Mad family of repressor proteins.

KLF11, also known as FKLF and TIEG2, is a repressor protein that reportedly recruits the mSin3A/HDAC complex through a short α-helical repression motif (α-HRM), termed SID as well, to mediate transcriptional silencing [5]. This SID is conserved in at least five KLF proteins (KLF9, 10, 11, 13, and 16) [5]. The KLF SID is sufficient to mediate interactions with mSin3A and transcriptional repression. Although both SIDs of KLF and Mad1 repressor proteins have the AA/VXXL core consensus and similar propensity for helix formation, the two SIDs can be classified into two subtypes on the basis of sequence similarity; in particular, the residues outside the AA/VXXL core sequence from PAH2 interacting consensus defined by Brubaker et al. [3,5]. Like Mad1 SID, KLF11 SID interacts with the PAH2 domain of mSin3A, which is necessary for the cell growth regulatory activity of the KLF protein. Therefore, it is logical to propose that the interaction with the PAH2 domain of mSin3A may play an important role in cell growth regulation.

The existence of many SIDs raises a question of whether similar domains from different families of repressor proteins bind to PAH2 with the same or different binding modes. This question is important in understanding functions of a wide range of repressor proteins that interact with mSin3A. The NMR or X-ray structure of SID of KLF11 complexed with the mSin3A-PAH2 domain is currently not available. A theoretical three-dimensional (3D) structure of such a complex is desirable with regard to interactions of the KLF11–mSin3A complex at the atomic resolution, particularly, the interactions involved in the binding of SID. In this article, we report a 3D model of the mSin3A–PAH2 domain complexed with KLF11 SID that was determined by a combination of computational and experimental methods. The analysis of this model reveals structural differences between the binding of KLF SID and that of Mad1 SID. This finding offers a new insight of cell growth regulation and the structural basis for developing selective modulators of repressor proteins.

*Corresponding author. Fax: (1)-507-284 9111.
**Also corresponding author. Fax: (1)-506-255 6318.
E-mail addresses: pang@mayo.edu (Y.-P. Pang), urrutia.raul@mayo.edu (R. Urrutia).

0014-5793/03/$22.00 © 2003 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00749-X

	23	27	30	35
KLF11 SID (target)	L-E-**Q**-T-D-M-**E**-**A**-V-E-A-**L**-V			
	8	12	15	20
Mad1 SID (template)	N-I-**Q**-M-L-L-**E**-**A**-A-D-Y-**L**-E			

Brubaker motif ϕ-Z-Z-ϕ-ϕ-X-A-A-X-X-ϕ-E

where, X = nonproline, ϕ = hydrophobic, and Z = hydrophobic/charged

Fig. 1. Primary sequences for the Sin3 interaction domain of KLF11 and Mad1 and the Brubaker interaction motif.

2. Materials and methods

2.1. Complex preparation

2.1.1. The mSin3A–KLF11 complex. The 3D structure of the SID of KLF11 complexed with the mSin3A–PAH2 domain was determined as follows: (i) determining the 3D structure of the SID of KLF11 by homology modeling using the SID of Mad1 protein taken from the mSin3A–Mad1 complex (PDB code: 1G1E) [3] as a template (for primary sequences, see Fig. 1); (ii) obtaining the 3D structure of the mSin3A–PAH2 domain directly obtained from the mSin3A–Mad1 complex; and (iii) determining the 3D complex structure of the SID of KLF11 bound with the mSin3A–PAH2 domain by manually docking the SID of KLF11 to the mSin3A–PAH2 domain to achieve maximal intermolecular interactions between the two partners. The resulting complex was refined by a 2.0 ns (1 fs time step) molecular dynamics (MD) simulation (see below) after the following modifications: (i) addition of hydrogen atoms; (ii) protonation or deprotonation of the Arg, Lys, Asp, Glu and His residues; (iii) addition of counter ions to neutralize the charged residues; and (iv) energy minimizations of the added hydrogen atoms and counter ions. The protonation states of all ionizable residues (Arg, Lys, Asp, Glu and His) were determined by visual inspection. Arg and Lys residues were protonated, unless they were located in a hydrophobic environment. One Cl^- anion was introduced next to a protonated, cationic residue if this residue was more than 8 Å away from an anionic residue. Asp and Glu residues were deprotonated, unless they were surrounded by hydrophobic residues. One Na^+ cation was placed in the vicinity of a deprotonated, anionic residue if this residue was more than 8 Å away from a cationic residue. His residue was protonated if it was less than 8 Å away from an unbalanced acidic residue, otherwise it was treated as a neutral residue. The location of every counter ion was determined by an energy minimization with a positional constraint applied to all atoms of the system except for the counter ion.

2.1.2. The mSin3A–Mad1 complex. The initial structure used in the MD simulation of the mSin3A–Mad1 complex was taken from the corresponding NMR structure [3] and modified with the same procedures as used for the mSin3A–KLF11 complex.

2.2. MD simulations

2.2.1. The mSin3A–KLF11 complex. The MD simulation of the mSin3A–KLF11 complex was performed by employing the SANDER module of the AMBER 5.0 program [6] with the all-atom force field by Cornell et al. [7]. The values of the keywords in uppercase letters used by the AMBER program were described in parentheses. The MD simulation used (i) the SHAKE procedure for all the covalent bonds of the system (NTC = 2 and NTF = 2) [8]; (ii) a time step of 1.0 fs (DT = 0.001); (iii) a dielectric constant $\varepsilon = 1.0$ (IDIEL = 1.0); (iv) the Berendsen coupling algorithm (NTT = 1) [9]; (v) the Particle Mesh Ewald method [10] used to calculate the electrostatic interactions (BOXX = 63.2826, BOXY = 50.2977, BOXZ = 49.6926, ALPAH = BETA = GAMMA = 90.0, NFFTX = 49, NFFTY = 49, NFFTZ = 49, SPLINE_ORDER = 4, ISCHARGED = 0, EXACT_EWALD = 0, DSUM_TOL = 0.00001); (vi) the non-bonded atom pair list updated at every 20 steps (NSNB = 20); and (vii) default values for all other keywords not mentioned here. The SID of KLF11 complexed with the mSin3A–PAH2 domain was simulated in a TIP3P water box [11] with a periodic boundary condition at constant temperature and pressure (NCUBE = 20, QH = 0.4170, DISO = 2.20, DISH = 2.00, CUTX = CUTY = CUTZ = 8.2, NTB = 2, TEMP0 = 298, PRES0 = 1.0 and NTP = 1). The resulting system (14 435 atoms) was first energy-minimized for 500 steps to remove close van der Waals contacts, and then slowly heated to 298 K (10 K/ps). A weak harmonic restraint in the Cartesian space (NTR = 1 and the harmonic potential force constant = 0.01 kcal mol^{-1}) was applied to the added counter ions to avoid large separations of these ions from the protein during the 2.0 ns MD simulation.

2.2.2. The mSin3A–Mad1 complex. The MD simulation of the mSin3A–Mad1 complex was carried out using the same method as described above. The mSin3A–Mad1 complex solvated by the TIP3P water molecules had 16 529 atoms. The Particle Mesh Ewald method accordingly used the following parameters: BOXX = 66.3087, BOXY = 57.7339, BOXZ = 47.6234, ALPAH = BETA = GAMMA = 90.0, NFFTX = 49, NFFTY = 49, NFFTZ = 49, SPLINE_ORDER = 4, ISCHARGED = 0, EXACT_EWALD = 0, and DSUM_TOL = 0.00001. The MD simulation was carried out for 2.0 ns (1 fs time step).

2.3. Structural analysis

The CARNAL module of the AMBER 5.0 program was used to calculate the time-average structure from the instantaneous structures obtained at 1 ps intervals during the entire 2.0 ns simulation. The root mean square deviation (RMSD) between the two structures of interest

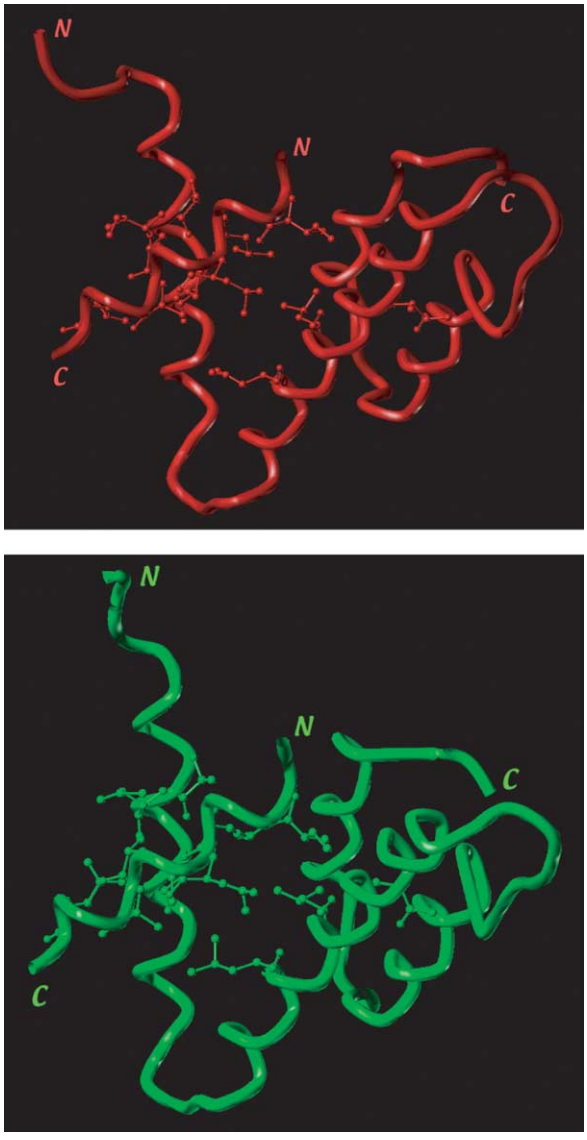


Fig. 2. Stereo view of backbone models of the average structure of the mSin3A–Mad1 complex derived from the MD simulation (top panel) and the NMR spectroscopic analysis (bottom panel). The hydrophobic side chains at the peptide–protein interface are shown with the ball-and-stick model.

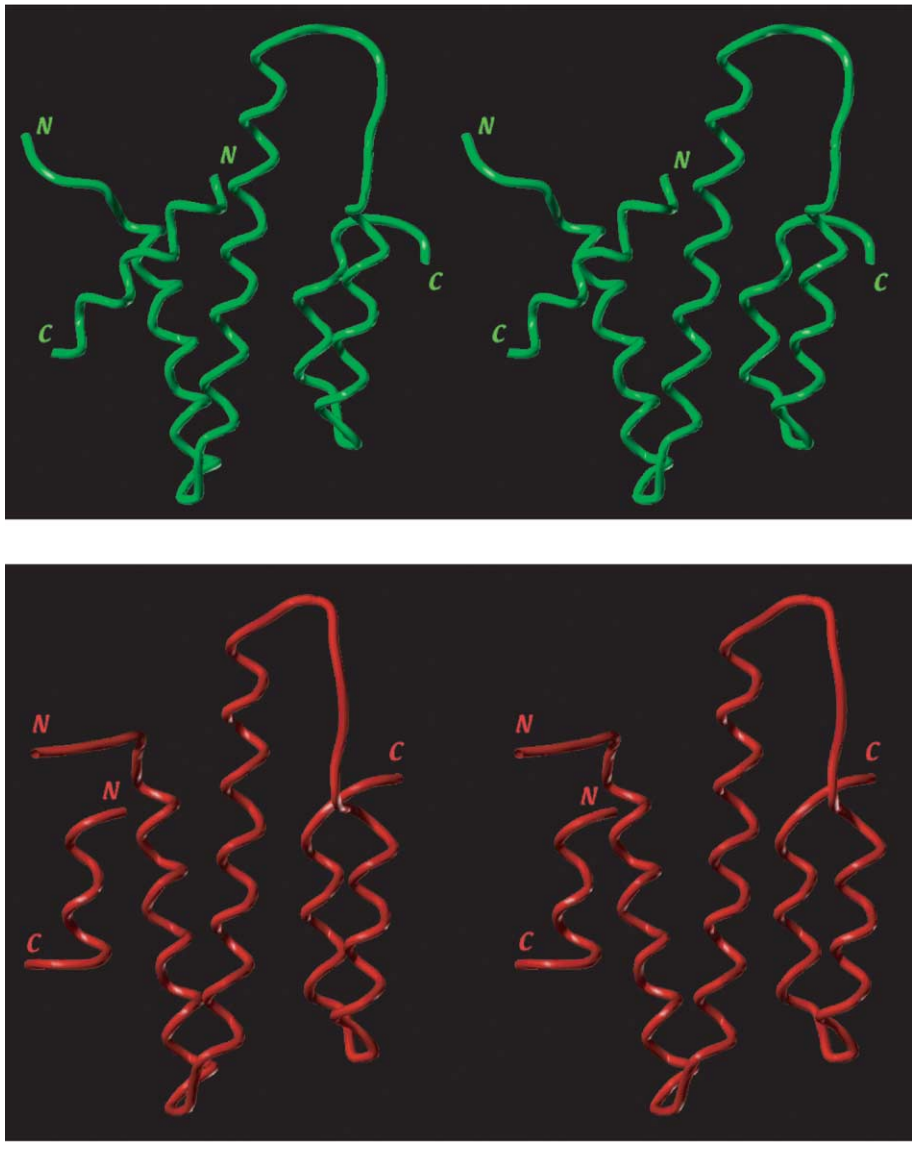


Fig. 3. Stereo view of backbone models of the SID of Mad1 complexed with the mSin3A–PAH2 domain (top panel) and the SID of KLF11 complexed with the mSin3A–PAH2 domain (bottom panel).

was calculated by using the molecular similarity function available in the QUANTA 97 program (San Diego, CA, USA).

2.4. Site-directed mutagenesis

2.4.1. Plasmid constructs. The nucleotide sequence of PAH2 domain-flanking amino acid residues 275–455 was obtained from pCMV2B-PAH2 [5] and subcloned in frame with GST into the pGEX-5X-1 expression vector (Pharmacia, Piscataway, NJ, USA). The wild type SID of KLF11 and its double-proline mutant, Dm (ΔE29P, ΔA30P), were obtained from GST expression vectors (GST-R1/R1m) as previously described [5] and subcloned in frame with Flag tag into the pCMV-Tag2 vector (Stratagene, La Jolla, CA, USA). Standard PCR-based methods were used to generate other mutant constructs (m1–4) within the KLF11 SID domain, including m1(ΔL23A), m2(ΔE29K), m3(ΔV31A), and m4(ΔE32K).

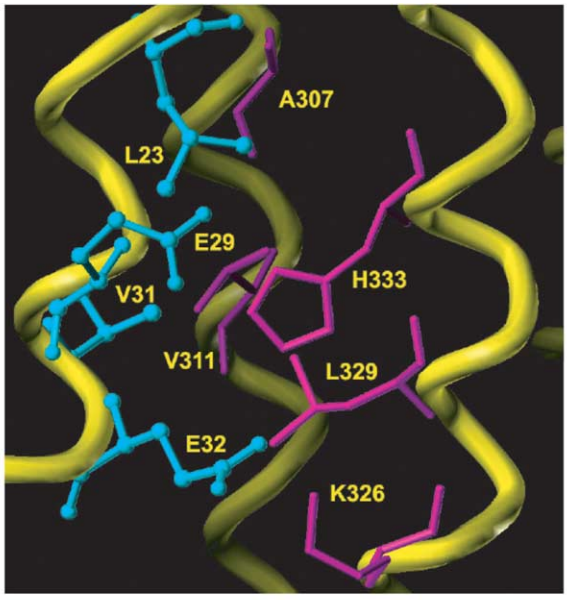
2.4.2. GST pulldown assays. GST–PAH2 fusion protein expression was induced in BL21 cells (Stratagene) by the addition of 1 mM isopropyl-1-thio-β-D-galactopyranoside for 3 h. Cells were lysed and subsequently purified using glutathione Sepharose 4B affinity chromatography according to the manufacturer's instructions. The wild-type SID and its various mutant proteins were produced by in vitro translation using the TNT-coupled transcriptional/translation system and T3 RNA polymerase in the presence of [^{35}S]methionine (Amersham Pharmacia BioTech, Piscataway, NJ, USA) according to the manufacturer's instructions. For the pulldown assays, the purified GST–PAH2 fusion protein (2 μg) bound to the glutathione-conjugated Sepharose beads was incubated with [^{35}S]-labeled Flag-R1 and its various mutants, respectively, in a lysis buffer (150 mM NaCl, 0.5% Nonidet P-40, 50 mM Tris–HCl, pH 7.5, 20 mM $MgCl_2$). Samples were incubated at 4°C for 3 h and then washed five times using the same buffer. The bound proteins were separated on 16.5% SDS–polyacrylamide gels. The gels were treated with AutoFluor (National Diagnostics, Atlanta, GA, USA), dried, and analyzed using a phosphoimager (Storm 860, Molecular Dynamics).

3. Results and discussion

An MD simulation of the mSin3A–Mad1 complex was first carried out as a control study to confirm that the time-average structure of the MD simulation is consistent with the average structure obtained from the NMR spectroscopic analysis. As apparent from Fig. 2, the time-average structure derived from a 2.0 ns (1.0 fs time step) MD simulation overlays well with the average structure obtained from the NMR spectroscopic analysis. The RMSDs of the backbone structure in the

mSin3A–Mad1 complex, mSin3A and Mad1, between the time-average MD structure and the average structure of the 15 NMR structures are 2.37, 2.43, and 1.94 Å, respectively. The intermolecular interactions involving hydrophobic residues between mSin3A and Mad1 in the computational and experimental structures are the same as shown in Fig. 2. In particular, the following hydrophobic interactions were observed in both NMR and MD structures: Ile9 of Mad1 interacts with Gln336, Val358 and Leu380 of mSin3A; Leu12 of Mad1 interacts with Phe379, Phe376, Phe328, Leu332, Ala307 and Val311 of mSin3A; Leu13 of Mad1 interacts with Leu332 and Leu329 of mSin3A; Ala15 of Mad1 interacts with Val311 and Ile308 of mSin3A; Ala16 of Mad1 interacts with Leu329 and Val311 of mSin3A; and lastly, Leu19 of Mad1 interacts with Ile308 and Val311 of mSin3A. These results validate the MD simulation method employed in the present study for the refinement of the mSin3A–KLF11 complex described below.

The 3D structure of the mSin3A–KLF11 complex derived from homology modeling and refined by a 2.0 ns (1.0 fs time step) MD simulation is depicted in Fig. 3. Similar to the mSin3A–Mad1 complex, the KLF11 complex consists of a four α-helix bundle for the PAH2 domain of mSin3A and an amphipathic α-helix for the SID of KLF11. Interestingly, the SID of KLF11 interacts with mSin3A in a different orientation relative to the mSin3A–Mad1 complex (Fig. 3). A 2.0 ns MD simulation reveals that the SID of KLF11 prefers to interact with mSin3A in a different orientation with its α-helix



The interactions between hydrophobic residues

LEU23 (SID) and VAL311 (PAH2) & ALA307 (PAH2)

VAL31 (SID) and VAL311 (PAH2) & LEU329 (PAH2)

The interactions between charged residues

GLU29 (SID) and HIS333 (PAH2)

GLU32 (SID) and LYS326 (PAH2) & HIS333 (PAH2)

Fig. 4. The interacting residues of the SID of KLF11 (cyan) and the PAH2 of Sin3A (magenta) identified for the mutation studies of the mSin3A–KLF11 complex.

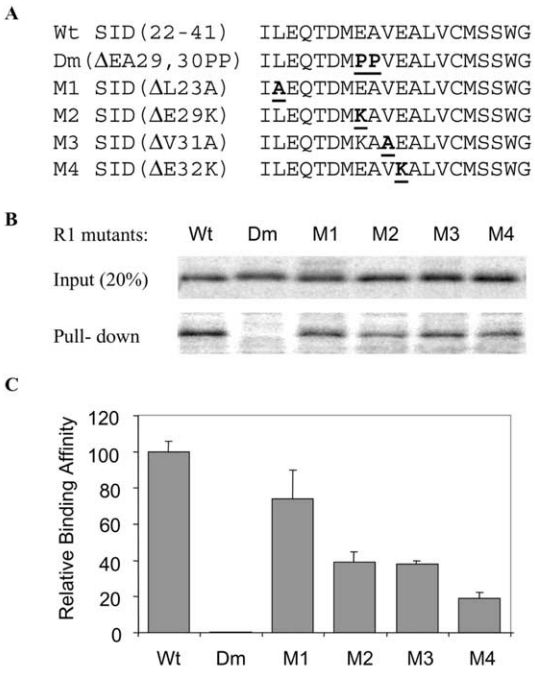


Fig. 5. Binding of KLF11 SID mutants to the PAH2 domain of Sin3a: In vitro translated Sin3a PAH2 was incubated with the corresponding GST SID mutants for binding assays. A: The sequence of the KLF11 wild type and the different SID mutants. B: Results of representative in vitro binding experiments showing the relative binding ability of the wild-type KLF11 SID and mutants to the Sin3a PAH2 (GST pulldown assays). C: Histogram showing the relative density of the PAH2 band resulting from the in vitro binding assays depicted in panel B.

parallel to helix 2 of the mSin3A–PAH2 domain, as shown in Fig. 3, although SID of KLF11 was initially positioned, just like the SID of Mad1 positioned in mSin3A, at an angle in an inter-helical buried position between the helices α-1 and α-2 of the PAH2 domain of mSin3A in the MD simulation. The hydrophobic interactions between KLF11 and mSin3A include: (i) Leu23 of KLF11 with Ala307 and Ile308 of mSin3A; (ii) Met28 of KLF11 with Leu332 and Gln336 of mSin3A; and (iii) Val31 of KLF11 with Val311 and Leu329 of mSin3A. The N-terminus (Ser295-Val302) of mSin3A in the mSin3A–KLF11 complex moves towards the KLF helix occupying the region where the N-terminus of the Mad1 helix was located in the Mad1 complex, although the MD simulations of the mSin3A–Mad1 and mSin3A–KLF11 complexes used the same setup. It is conceivable that a different binding mode for SID of KLF11 results from its sequence difference relative to SID of Mad1.

Site-directed mutagenesis studies were then carried out to validate the computational studies described above. Visual inspection of the mSin3A–KLF11 complex suggests that Leu23, Val31, Glu29 and Glu32 of SID interact with Val311, Leu329, His333 and Lys326 of PAH2, respectively (Fig. 4 and Table 1). According to the theoretical model of the mSin3A–KLF11 complex, (i) a mutation of Leu23 (SID) or Val31 (SID) to Ala is expected to reduce or diminish the hydrophobic interaction with Val311 (PAH2) and Leu329 (PAH2) and (ii) a mutation of Glu29 (SID) or Glu32 (SID) to Lys is expected to reduce or diminish the electrostatic interaction with His333 (PAH2) and Lys326 (PAH2). The

Table 1
Site-directed mutation analysis of the SID of KLF11

Mutation of residues in the SID of KLF11			Interacting residues in the PAH2 of mSin3A
From	ID	To	
Leu	23	Ala	Val311, Ala307
Glu	29	Lys	His333
Val	31	Ala	Val311, Leu329
Glu	32	Lys	Lys326, His333

Glu29Pro/Ala30Pro double mutant (Dm) was used as a negative control, as this mutant does not interact with the PAH2 domain of Sin3A [5]. The sequences containing these mutations are listed in Fig. 5A.

To measure and compare the relative binding affinity of KLF11 SID and its various mutants for their purported interaction with the PAH2 domain of Sin3A, GST pulldown assays were performed by using GST–PAH2 and [^{35}S]methionine labeled in vitro translated SID and its mutants. A representative autoradiography of the pulldown SID and its mutants by GST–PAH2 is shown in Fig. 5B, and the results from three independent experiments are summarized in Fig. 5C.

Dm with two proline mutations within the KLF11 SID does not bind PAH2 and is consistent with our previous report [5]. As expected, the mutant constructs M1–M4 defined in Fig. 5A bind to the Sin3A–PAH2 domain with lower affinities than the wild-type SID domain; a mutation of Leu23 (SID) or Val31 (SID) to Ala reduces the binding affinity of SID of KLF11 to 75% and 40% of the wild type, respectively. Likewise, a mutation of Glu29 (SID) or Glu32 (SID) to Lys decreases the binding affinity of SID of KLF11 to 40% and 20% of the wild type, respectively (Fig. 5C). These experimental results are consistent with the predicted interactions between the SID residues of KLF11 and the PAH2 domain residues of Sin3A as listed in Table 1, thus supporting the theoretical model of the KLF11–Sin3A complex.

At least three types of DNA-binding transcription factors (i.e. the Mad proteins, KLF repressor proteins and the yeast transcription factor Ume6) have been shown to interact with mSin3A through the PAH2 domain. A recently identified corepressor protein Pfi has also been shown to interact with PAH2, although it lacks the DNA-binding ability [12]. Interestingly, all these proteins bind PAH2 through an α-helical motif. Although these SID motifs are structurally similar, their binding affinities for PAH2 are different. The SID of the Mad1 (29 nM [3]) binds PAH2 with higher affinity than the SID from KLF11 (90 nM; manuscript in preparation). Pf1 and Ume6 bind with much lower affinities [12,13]. The presence of a PAH2-interacting SID among different families of transcription factors poses the question of whether the interactions of these proteins with the same PAH2 domain of Sin3A are the same or different. An answer to this question helps to understand the function and regulation of these interactions. In this study, we have developed the 3D structure of the Sin3A–PAH2 domain complexed with a 13-residue fragment of SID of KLF11 using homology modeling, MD simulation and site-directed mutagenesis. This theoretical model reveals that the binding of KLF11 SID in its Sin3A complex is different from that of Mad1 SID in its mSin3A complex.

The first PAH domain of Sin3 shares approximately 40% sequence similarity with PAH2, and thus has the same folding as PAH2 (Urrutia et al., unpublished data). Although several PAH1-interacting repressor domains have been identified (e.g. Rest), the intermolecular interactions of these domains are not known. Based upon the results described above, we predict that similar sequences with a propensity to form the α-helical motif bind the hydrophobic pocket formed by the folded PAH1 domain in different orientations. Such variations may provide selectivity and differential affinity of different repressor proteins for Sin3A for a systematic regulation of cell growth.

Accession Number: The mSin3A–KLF11 complex structure is available at the PDB (PDB code: 1PO4).

Acknowledgements: Supported by the DARPA (DAAD19-01-1-0322 to Y.-P.P.), the University of Minnesota Supercomputing Institute (Y.-P.P.), the NIH (DK52913 and DK56620 to R.U.), the Lustgarten Foundation for Pancreatic Cancer Research (to R.U.), and the Thompson-Mayo Postdoctoral Fellowship (to G.A.K.).

References

[1] Ayer, D., Lawrence, Q. and Eisenman, R. (1995) Cell 80, 767–776.
[2] Eilers, A., Billin, A., Liu, J. and Ayer, D. (1999) J. Biol. Chem. 274, 32750–32756.
[3] Brubaker, K., Cowley, S., Huang, K., Loo, L., Yochum, G., Ayer, D., Eisenman, R. and Radhakrishnan, I. (2000) Cell 103, 655–665.
[4] Spronk, C., Tessari, M., Kaan, A., Jansen, J., Vermeulen, M., Stunnenberg, H. and Vuister, G. (2000) Nat. Struct. Biol. 7, 1100–1104.
[5] Zhang, J., Moncrieffe, M., Kaczynski, J., Ellenrieder, V., Prendergast, F. and Urrutia, R. (2001) Mol. Cell. Biol. 21, 5041–5049.
[6] Pearlman, D.A. et al. (1995) Comput. Phys. Commun. 91, 1–41.
[7] Cornell, W.D. et al. (1995) J. Am. Chem. Soc. 117, 5179–5197.
[8] Ryckaert, J.P., Ciccotti, G. and Berendsen, H.J.C. (1977) J. Comput. Phys. 23, 327–341.
[9] Berendsen, H.J.C., Postma, J.P.M., van Gunsteren, W.F., Di Nola, A. and Haak, J.R. (1984) J. Chem. Phys. 81, 3684–3690.
[10] Darden, T.A., York, D.M. and Pedersen, L.G. (1993) J. Chem. Phys. 98, 10089–10092.
[11] Jorgensen, W.L., Chandreskhar, J., Madura, J.D., Impey, R.W. and Klein, M.L. (1982) J. Chem. Phys. 79, 926–935.
[12] Yochum, G. and Ayer, D. (2001) Mol. Cell. Biol. 21, 4110–4118.
[13] Washburn, B. and Esposito, R. (2001) Mol. Cell. Biol. 21, 2057–2069.